

STRUCTURED EXIT INTERVIEWS USING MDS

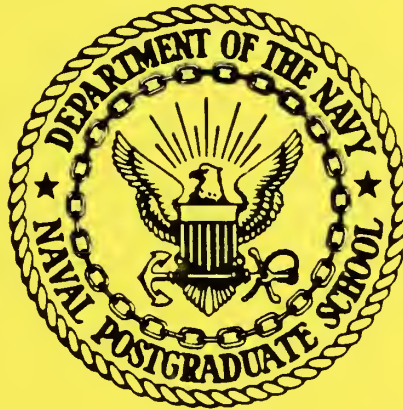
by

Robert R. Read

NPS55-85-024

NAVAL POSTGRADUATE SCHOOL

Monterey, California



TECHNICAL

STRUCTURED EXIT INTERVIEWS USING MDS

By

Robert R. Read

September 1985

Approved for public release; distribution unlimited

Prepared for:
Naval Postgraduate School
Monterey, CA 93943-5100

B F
39.5
R 42
1985
C. 2.

NAVAL POSTGRADUATE SCHOOL
Monterey, California

Rear Admiral R. H. Shumaker
Superintendent

David A. Schradly
Provost

Reproduction of all or part of this report is authorized.

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM	
1. REPORT NUMBER NPS55-85-024		2. GOVT ACCESSION NO.	
4. TITLE (and Subtitle) STRUCTURED EXIT INTERVIEWS USING MDS		3. RECIPIENT'S CATALOG NUMBER	
		5. TYPE OF REPORT & PERIOD COVERED Technical	
		6. PERFORMING ORG. REPORT NUMBER	
7. AUTHOR(s) Robert R. Read		8. CONTRACT OR GRANT NUMBER(s)	
9. PERFORMING ORGANIZATION NAME AND ADDRESS Naval Postgraduate School Monterey, CA 93943-5100		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS	
11. CONTROLLING OFFICE NAME AND ADDRESS Naval Postgraduate School Monterey, CA 93943-5100		12. REPORT DATE September 1985	
		13. NUMBER OF PAGES 43	
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified	
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE	
16. DISTRIBUTION STATEMENT (of this Report) Approved for public release; distribution unlimited.			
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)			
18. SUPPLEMENTARY NOTES			
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) MDS, Multidimensional Scaling, Measurement of Teacher Performance			
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) The use of multidimensional scaling (MDS) is explored as a tool for discovering important dimensions of student group perception of instruction. Proximity data were collected at computer terminals with the aid of user friendly programs. The technique allows the dynamic discovery of characteristics of instruction that are most important to the particular student group. Unwanted effects in the scoring of instructors can be removed without the requirement of a cross classified experimental design. Some specialized methods for displaying the output of the KYST program are introduced.			

I. Introduction.

The main purpose of this report is to introduce the technique of MDS (Multidimensional Scaling) as a tool for organizing, enhancing, and structuring information that may be obtained from students during their exit interviews. More specifically we are concerned with the question of measuring and summarizing the students' perception of the instructional treatment they received while at NPS. The administration is obliged to monitor this process and MDS offers a dynamic and yet structured way to manage this problem. Moreover, it will be seen that the technique is a subtle one which allows the discovery of new factors that influence the perception process. It has the potential of providing a way to separate unwanted effects.

Recent advances in computer input technology make feasible the data collection component that is inherent in the application of the MDS technique. The student may link to a user friendly computer program which will request information of the proper kind. Responses are input by moving the cursor to the proper position and striking an appropriate key. (The use of a touchscreen or a mouse would be even better.) When finished, the respondent can send his input to a central file where it is merged with input from other sources and processed. The use of the console for the administration of a questionnaire allows much information to be gathered in a reasonably short period of time. The type of information requested and the way it is analyzed are the main issues treated herein.

A secondary but useful aspect of this report is to review the history of student-instructional information collection here at NPS. This is done in Section II. Readers who are uninterested in history may proceed directly to Section III which contains a description of the MDS technique as applied to some developmental work performed with the graduating students in the

Operations Analysis Curriculum. The results of this work are analysed and summarized in Section IV (which also contains material comparing different information display techniques). Some data documentation is included in the appendices. The remainder of this introduction is devoted to mentioning some shortcomings of the SOF system currently in use. The usefulness of MDS as supplementary and enhancing the SOF will become apparent.

The measurement of teaching quality at NPS in recent years has been largely through the use of the data summaries obtained from the SOF system. Although a number of weaknesses of this administrative use have been identified, there is little tangible evidence that any other information is being used as well. Sources of supplemental information might include some type of systematic review of the course journals, and some method for measuring how much the students have learned or how much they have grown as students.

It appears that resources for measuring student progress will not become available. Occasionally classroom visitation has been mentioned as a source of information. It certainly can be valuable for instructor development, but the potential for abuse is great and it may be damaging if used for measurement.

Experience with the SOF and similar systems has not been satisfactory.

Problem areas include:

1. Data collected in one quarter in conjunction with data collected at other times or involving other students or both are used to make cross comparisons between instructors.
2. The SOF data is not collected under controlled experimental conditions.
3. The set of instructor rating scales is static.
4. It is limited to the students' perception of instruction.

In addition, no provision is made for determining whether the instructor covered the correct material and in sufficient depth. The presence of SOF and its perceived use can have a subtle and corrupting effect. It encourages instructors to compromise when choosing between what is right and what is popular.

As of this writing the practice of interviewing students as they exit is not institutionalized. Each department or curricular office utilizes this opportunity as they see fit. Since graduation can be a very busy time for the student, it is recommended that any organized information effort take place early in the last quarter of instruction. We cite the following items in relations to such a system.

Advantages: All the information is collected at the immediate end of the educational experience. The system envisioned allows for the dynamic discovery of factors of instruction that are of import to each individual class. In addition to the development of instructor rating scales, the "treatment" given to each class is summarized. Such collateral information could have value for curricular development and for scheduling.

Disadvantages: The students may have difficulty comparing instructors they have seen in the distant past with those they have seen recently.

II. History.

The collection of student-instructor evaluation information has a spotty history prior to 1972. Many department chairmen held informal "exit interviews" with graduating students. Some departments developed questionnaire forms which could be used by their faculty at individual option. About this time it became popular for institutions to use SIR (i.e. the Student-Instructional Report developed and processed by the Educational Testing Service at Princeton). SIR is a thirty-nine item questionnaire to be filled out by each student in each

course and sent to Princeton for processing. It was used here at NPS a few times in response to mounting pressure to have a uniform school-wide policy in this area.

Because of the expense, the length, and the large return-time involved with the adoption of SIR, the Faculty Council formed a committee to consider the development of a shorter form that was more appropriate for our needs and which could be processed locally. Support was made available and development took place. Much of the details of this activity is reported in the joint master's thesis of Burgess and Vaughn. Using the results of this thesis the committee developed the SOF, which has been in use ever since.

In a 1972 study, Read and Zweig explored the effects of using several different scoring methods applied to the same set of student survey data. An important result was that, from the point of view of the instructors, the choice of the scoring method can lead to some rather sharp differences in their rankings. Other results of this paper indicate that; i) data of this type cannot discriminate well among the non extreme teachers, and ii) there is difficulty in collecting detailed information from students when that information is based on experiences over one year old.

In the work mentioned earlier, Burgess and Vaughn performed factor analysis studies on the large data sets collected from our graduating students under the auspices of the Faculty Council. The technique involved the specification of eighty-six binary discriminators. Each student marked whether or not each of his instructors at NPS possessed the attribute for each of the eighty-six items. He also rated the overall instructional quality of each of his instructors on the "ladder" scale introduced by Elster, et al. (see Read-Zweig). The collection of these scales enabled the identification of two subsets of instructors, good and poor, as perceived by our students.

Factor analysis studies were performed on the eighty-six dimensional space of scores restricted to the union of the two subsets. There were differences among the several curricula, but generally the resulting factor spaces were seven dimensional and the principal factors identified are:

- i. Organization and clarity.
- ii. Instructor individual interaction.
- iii. Evaluation technique.
- iv. Synthetic-analytic approach.
- v. Stimulation.
- vi. Dynamism and enthusiasm.
- vii. Instructor group interaction.

Based upon the studies contained in this thesis, the special committee of the Faculty Council developed the SOF form. There have been no important modifications since.

The first eleven items of SOF ask the respondent to indicate his level of agreement or disagreement, on a scale of one to five, to statements about behavioral characteristics which are sharpened versions of the seven factors listed above. The eleven items appear in section III below. The next five items request overall ratings of instructor, course, text, exams, and laboratories on a nominal, but ordered, scale also ranging from one to five. Additionally, there is provision for voluntary free form comments. These comments constitute private communication from the student to the instructor. The data are collected in the last week of instruction of each term. They are machine processed and returned in the third or fourth week of the next term. Some studies have been made of the SOF data and the results are reported below.

It has become popular to use the class average response on SOF Item 12 to rank the instructors within a given department. The appropriateness of this is questionable and, as a result, the author (see Reference 5) was authorized to do a specialized study on some data made available from the OA Curriculum. The particular data selected has an unusual advantage in that it can be cross classified. That is,

- i. It involved eight student groups whose personnel was stable for each of three successive courses in the probability and statistics sequence.
- ii. The 24 classes (i.e. 3×8) were taught by a set of seven instructors.

Thus it is feasible to analyse the responses to item 12 using a cross classified experimental design. With respect to the student groups and the courses the experimental design is balanced, but not so with respect class size and instructors. The class section sizes ranged from 13 to 47. No single instructor taught all three courses but five of the instructors taught two of the three. It was deemed fortunate to do this well.

The mean value for item 12 was modeled as the sum of a student group effect, a course effect, and an instructor effect. All other effects were included in the error term of a standard analysis of variance model. All three main effects were highly significant. The F statistic for instructors was about 20 standard deviations to the right of its mean; for courses about 50 standard deviations; and about 30 standard deviations for student groups. Thus the effect of the course is paramount and the effect of the student group is more important than the effect of the instructor.

The resulting change in instructor rankings was quite noticeable (see Reference 5). That is, the ranking based on average response to item 12 compared to the ranking produced by the instructor effect estimates are

different. The replacement of rankings of this latter type for those of the former would be useful if cross classified data could be found on a widespread basis. Generally they cannot.

In this same study a discriminant analysis was performed on the first 13 SOF items (but omitting number 12 which was used to define the groups) in an effort to learn if these items could be used to cluster the instructors according to their scores on item 12. These SOF scales can be identified as the last 13 scales listed in Table 3. The discriminant space was one dimensional (i.e. 98% of the total variance was contained in the first principal component) and its direction was dominated by item 13, the course rating. This result was consistent over three consecutive quarters of data. The three direction cosines (of item 13) hovered around one-half.

Recently, the author and one of his students have applied the MDS technique for purposes of discovering what is important to students. User friendly programs were developed by Lt. J. McCourt as part of his master's thesis work, and they were tested with the March 1985 graduating class in curriculum 360. Much was learned in the areas of data collection and interpretation, and a number of modifications are suggested. McCourt also applied techniques of regression, factor analysis, and cluster analysis. Apart from the results that appear later in this report, this thesis confirmed a number of earlier results. E.g. the factor space of the SOF data is still one dimensional as it was in 1976. Also course organization accounts for the largest share (see Burgess and Vaughn) of the total variability of a proposed MDS solution. The results of cluster analysis emerged as a most valuable tool in aiding the students' interpretation of their MDS perception spaces.

III. General Description of MDS

Consider a data matrix of the following form:

		FACTORS				
		F1	F2	F3	...	FK
INSTRUCTORS	1					
	2				.	
	3				.	
	X(i,j)	...
	.				.	
	N.				.	

The elements $X(i,j)$ represent the score given to the i th instructor for the j th factor. For example, the SOF system has this structure. The K factors are prescriptive in nature and 13 or so in number. (See the last 13 scales in Table 3.) Each student provides entries for each instructor that he has had for each quarter on a scale of one to five (which is treated as an interval scale). By focusing upon a single instructor for a given class, we have a distribution of scores for each factor. These may be summarized by using the median (say) for that instructor's line entry in the data matrix.

Now suppose that we do not want to be prescriptive about the factors that may appear in the score summary table, indeed we do not even want to choose K in advance. How might we generate such information, and having done so, how might we interpret it? The multidimensional scaling technique developed by the behavioral scientists provides an answer. Since the factors are unspecified, the information requested from the respondent must have an indirect form. Then it must be converted into the above data format.

Interpretation of the factors can be accomplished in a number of ways. For this purpose have employed the application of cluster analysis and multiple regression, and then follow-up interviews with the respondents.

Let us be more specific. The MDS approach to this kind of subtle and indirect data collection lies in asking the subjects (students) to provide "proximity data" for each pair of instructors that they have had. For a definite example, let us focus on a particular pair of instructors, say 1 and 4, and ask the student to rate how dissimilar they are on a scale of, say, one to nine; one meaning that their dissimilarity is very low (i.e. they are virtually identical) and nine meaning that they are as different as they can possibly be. Such information is called proximity data, but we need to be more explicit about the general nature of the discrimination. Thus, it would not do to allow a score of one simply because both 1 and 4 are very tall, or a score of nine because 1 weighs 120 pounds and 4 weighs 260 pounds. It is necessary to focus the dissimilarities to those general areas of teaching effectiveness that are important to the respondent (but it is not necessary for the respondent to be able to articulate just what these scales are).

The above point raises one of the first issues in designing experiments of this type. How should the proximity criterion be verbalized? Possibilities include:

- (i) Their ability to induce me to learn.
- (ii) Their general skill as instructors.
- (iii) My general educational enhancement as a result of having taken courses from them.
- (iv) Their general effectiveness as an instructor.

I believe that you'll agree that these are rather general scales, in fact they are multidimensional and our goal is to discover the number and nature of the individual dimensions that make up this composite.

To continue, let us suppose that each subject has provided dissimilarity data of this type for each pair of instructors that he has seen, and that the results have been summarized into a triangular matrix which may be likened to a mileage table that one finds on roadmaps. E.g. perhaps eighteen students have seen both instructors 1 and 4, and the median of these eighteen values have been entered into the table. It is a triangular table because the proximity of 1 and 4 is the same as the proximity of 4 and 1. There will be no entries on the diagonals.

The conversion of proximity data into the factor score data, $X(i,j)$, is accomplished using the MDS program KYST. More information about this technique is presented in References 2 and 3. For now, we need only be concerned with some remarks about how well the conversion can be done, and the uniqueness of the result.

Firstly, there are $N(N-1)/2$ values of proximity and $N \times K$ values in the factor score matrix. To obtain any kind of solution the former must be greater than the latter, so we must have

$$K \text{ less than } (N-1)/2.$$

In fact, experienced workers with this technique provide us with the thumb rule that K should not be more than 25% of N in order to get good results, (See Kruskal and Wish).

Secondly, we do not expect to get an exact conversion. The input numbers reflect the perceptions of the respondents, and such data cannot be expected to conform to rigid mechanical standards. The result will be a compromise much like the compromise made when a function is fitted to data using least squares.

Thirdly, the result can only be unique up to an orthonormal (i.e. distance preserving) transformation applied to the factor score matrix; the proximities only emulate distances between rows of this matrix.

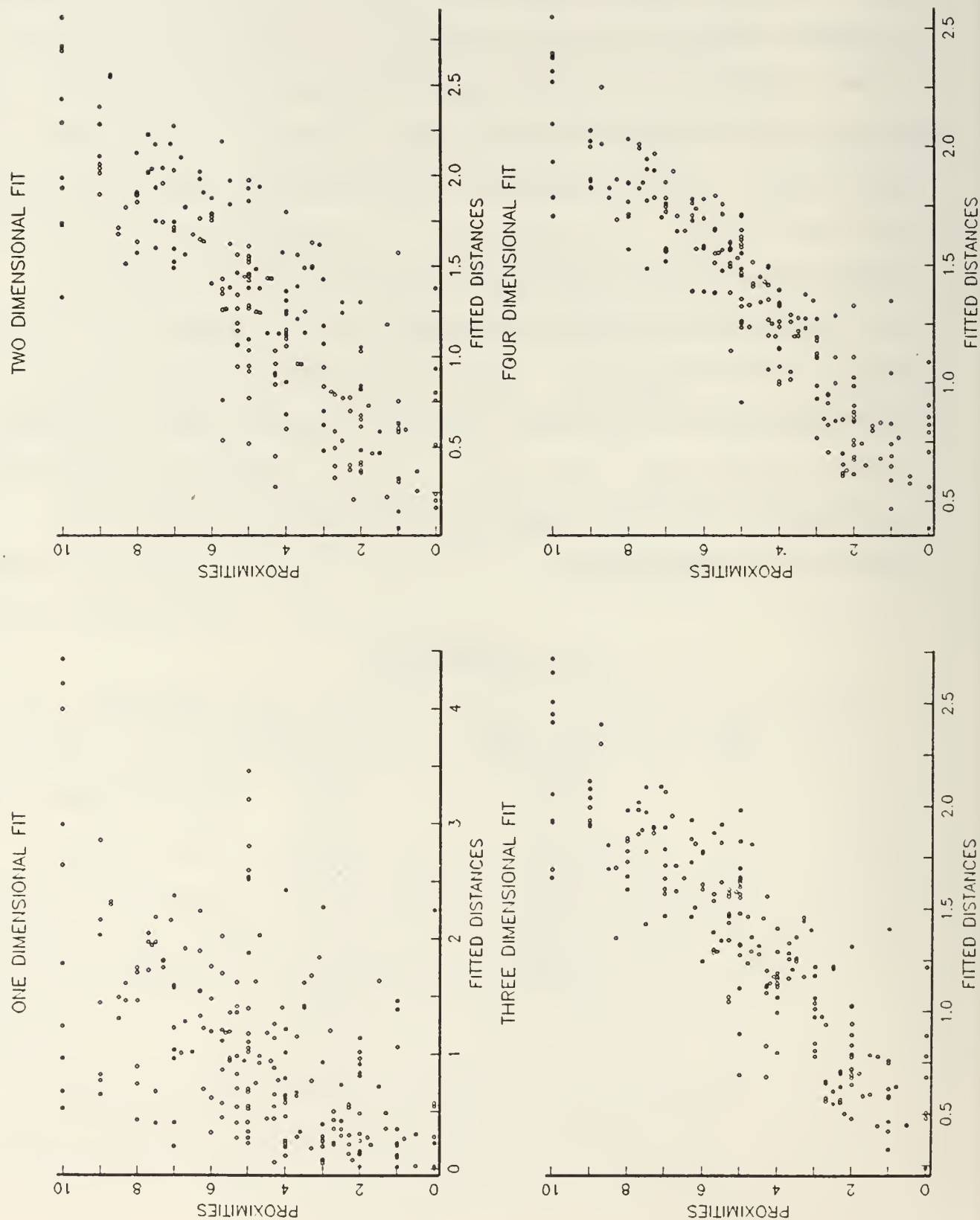
Choice of Dimension.

Typically the user of an MDS program will want to experiment with the value for K before he settles upon a final configuration. The goodness of fit is judgemental as there are no formal statistical tests to help decide. To this end a scatter plot of the input proximities against the computed row distances (i.e. from the fitted factor matrix) for each K is useful. See Figures 1 and 2. Further, one seeks a point of diminishing returns in a table or plot of stress vs K, where stress is the goodness of fit (i. e. a normalized sum of squared distances between proximities and corresponding factor distances for a K dimensional solution). Table 1 contains this information for the present work. These functions appear to decline at a rather uniform rate and, for our data, there is no obvious "knee" in the curve to use for the choice of K. Thus arbitrary choices will be made in the present example. Further discussion of this problem appears in the conclusions and recommendations.

Table 1
Stress vs. Dimension

Dim.	Sept. 1984	Mar. 1985
K	Stress	Stress
1	0.653	0.463
2	0.409	0.354
3	0.323	0.310
4	0.244	0.242
5	—	0.195
6	—	0.145

FIGURE 1
PROXIMITIES VS FITTED DISTANCES - SEPT 1984

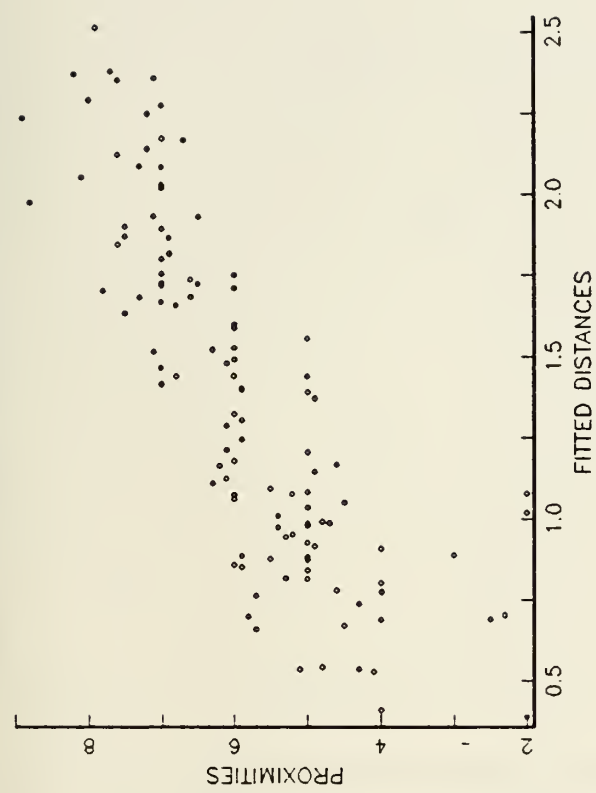


PROXIMITIES VS FITTED DISTANCE - MARCH 1985

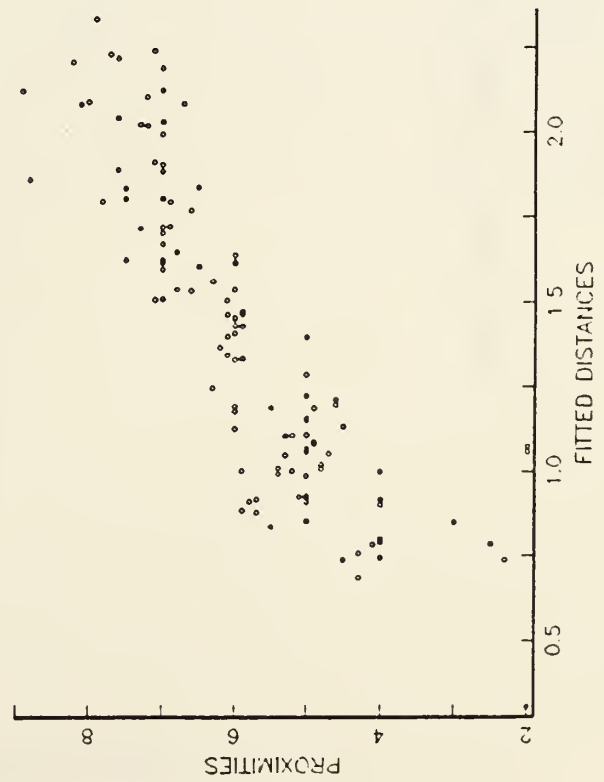
TWO DIMENSIONAL FIT



THREE DIMENSIONAL FIT



FOUR DIMENSIONAL FIT



FIVE DIMENSIONAL FIT

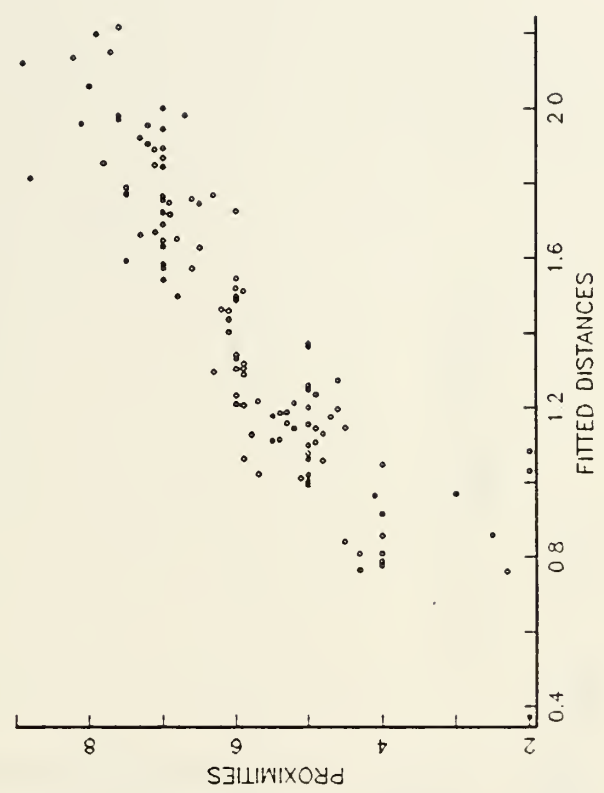
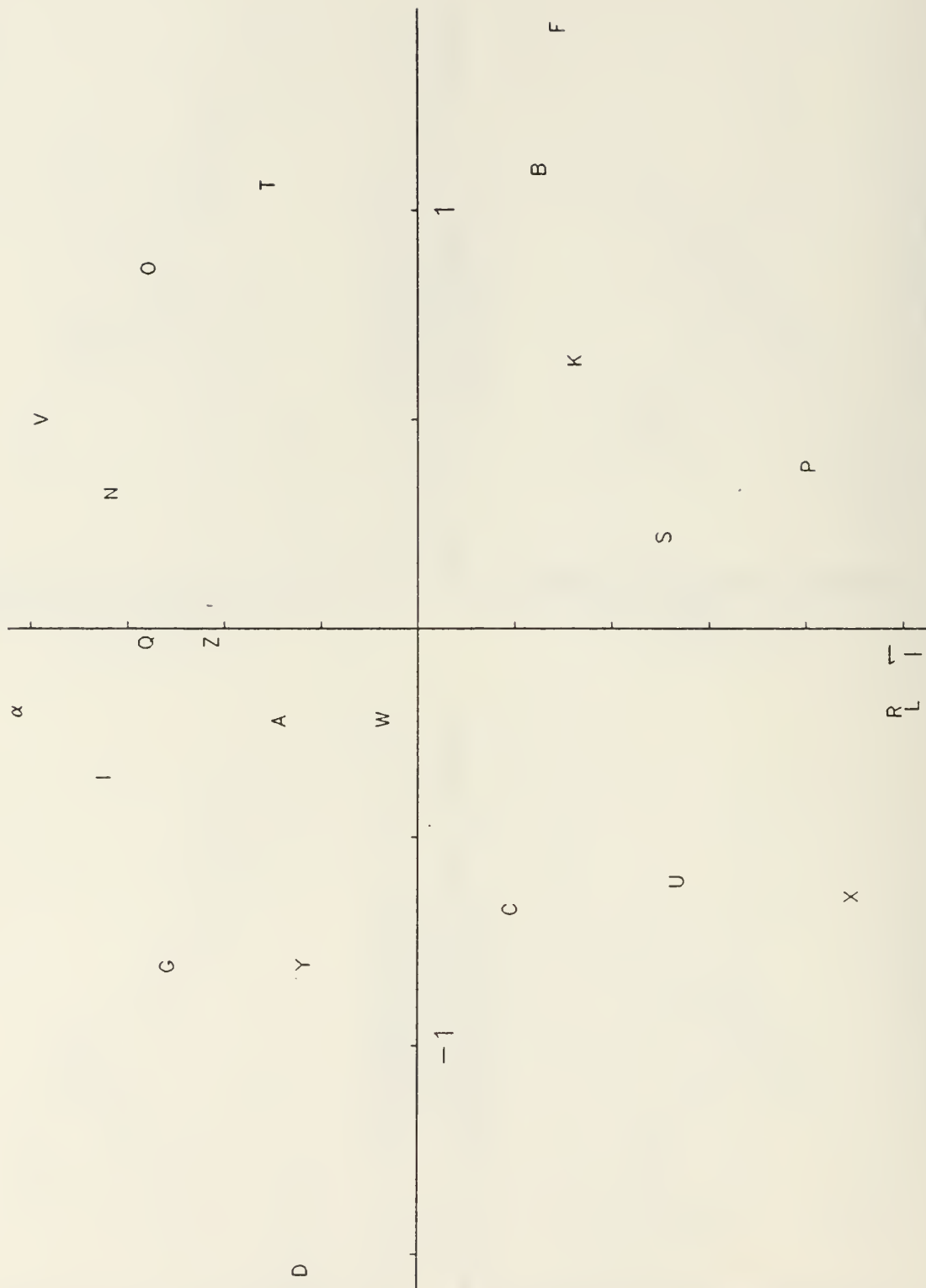


FIGURE 3

PRINCIPAL COMPONENTS - SEPT 1984 DATA



FIRST PRINCIPAL COMPONENT

SECOND PRINCIPAL COMPONENT

FIGURE 4

PRINCIPAL COMPONENTS - SEPT 1984 DATA

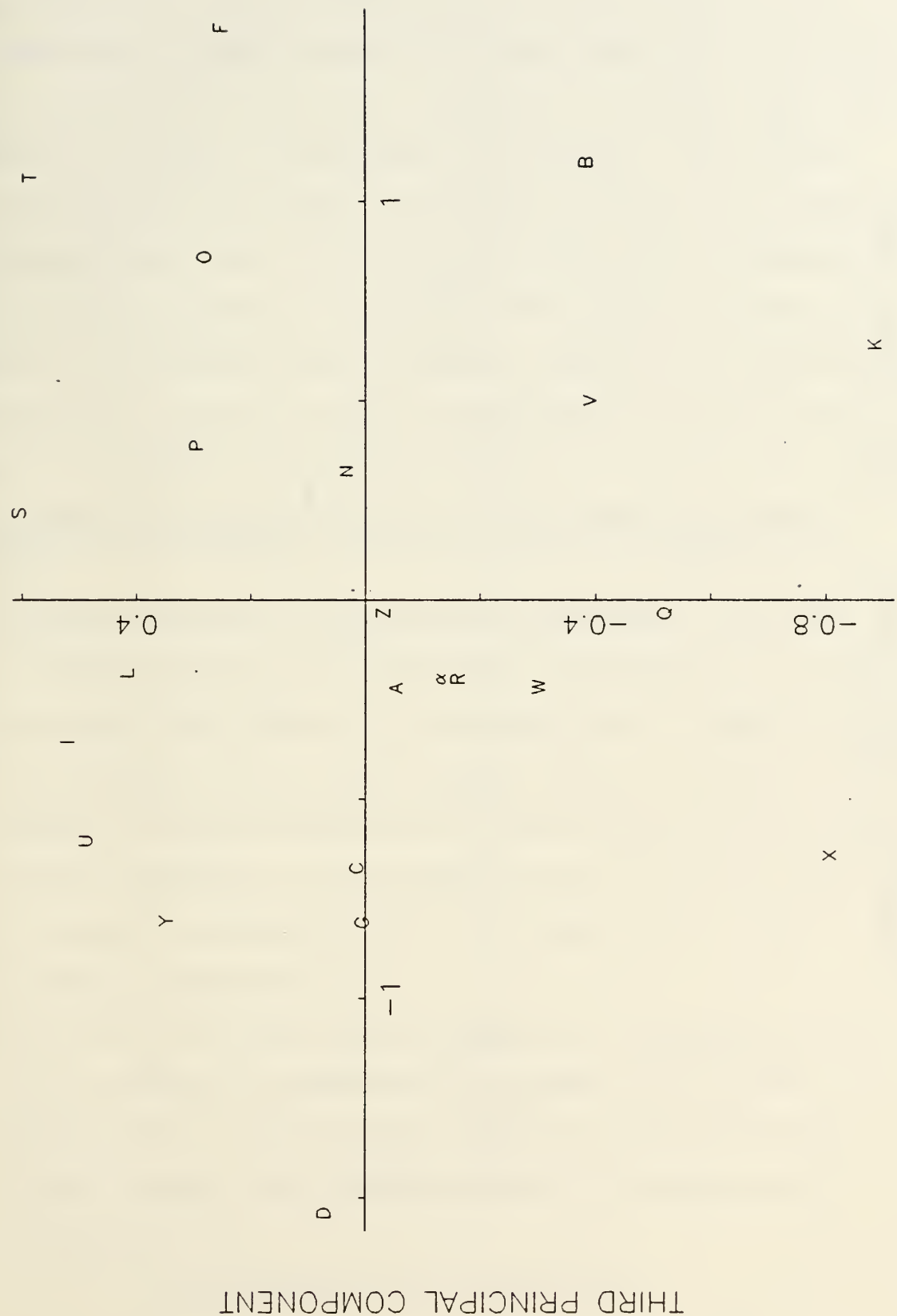


FIGURE 5

PRINCIPAL COMPONENTS - SEPT 1984 DATA



SECOND PRINCIPAL COMPONENT

Interpretation of factors.

Having settled upon K and recorded the results of the KYST program into our factor score matrix, our next step is to interpret the columns of this matrix. Often this is done by artwork i.e., scatter plots of $X(i, j)$ vs $X(i, k)$ for $i=1, \dots, N$ and for all fixed pairs of axes (j, k) are prepared. Each point is identified by the name of the instructor it represents. This provides a spatial representation of the objects (instructors); and the subjects, with the aid of the experimenter, often can identify the nature of the dimensions by viewing the relative positions of the objects. This stimulates thought for explanations of why they are placed as they are. Examples of this may be found in the references. Also, this is what the exit interview is about. But to be assured that the interview has greatest productivity, it is wise to be prepared with data summaries and graphical displays. We draw attention to three techniques:

1. Rotation of the MDS solution to principal components. Since the solution is invariant under orthonormal transformations, it is wise to rotate it to a form such that projections to the planes of pairs of coordinate axes will reveal as much structure as possible. The choice of principal components has proved useful and has become the default presentation. The literature in Factor Analysis can be consulted for alternative choices. We present some of our own later on in section IV.
2. Application of a Cluster Analysis algorithm. The identification of disjoint groups of instructors will facilitate in the discovery of instructional characteristics that are held in common within the several clusters. We use the K-means algorithm and some ad hoc techniques.

3. Multiple Regression of peripheral data on the MDS solution space. Our respondents were asked to rate the instructors on a number of bipolar scales, such as: 1) This instructor required much work outside of class. 2) This course was more theoretical than applied. 3) This instructor or this course was reputed to have a difficult grading policy. 4) This course relied heavily upon the prerequisites. These ratings too were on a scale of one to nine. In addition, the 13 SOF factors were retrieved to serve as bipolar scales. Each such scale serves as a response variable for a multiple regression upon the factor scores (i.e. the solution provided by the MDS program KYST). Whenever the multiple correlation coefficient is high for a bipolar scale, the direction of the associated scale in the factor space can aid in the interpretation of its axes.

IV. Results of Experimental Work

September 1984 study.

A brief pilot study was done on short notice in September of 1984. The cooperation of seven members of the graduating class was enlisted, and they were asked to provide proximity data for all of their instructors in OA courses on a typed sheet of paper. This layout contained a matrix of blanks with all of the instructors they might have had marked on the margins. Their task was to select the pairs that pertained to them and decide how far apart were the members of the pair in terms of the quality of instruction provided and their ability to motivate learning. Times of up to one hour were reported to perform this task. These data are recorded in Appendix A in the form of median responses.

By the time that we were able to produce output from the KYST program there were only three of the students still available for an exit interview. A three dimensional solution was selected and they were shown the projection of the points

on the planes formed from all pairs of the three principal components. These plots appear as Figures 3, 4, and 5, except that for the interviews the letters were replaced by nemonics that identify the instructors.

The students were shown these plots and asked for their interpretation of the spatial configuration of points. There was some difficulty in doing this, and the character of the axes was not clear cut. But in general terms we have the following: The instructors on the right side of the first dimension let the students get much of the material out of the textbook, while those on the left did not. Dimension two seemed related to structure with high structure close to the bottom of Figure 3. The third dimension may be related to usefulness of the course material, the greater toward the bottom of Figure 5. It may be noted from the scale markings that the second principal component is not much smaller than the first, but the third is noticeably smaller. I.e. the data swarm in 3-space is rather flat.

On the other hand, the students deemed it easier to comment on the characteristics of clusters of instructors. Most prominent were the points in the fourth quadrant of Figure 3. These instructors (specifically B, K, S and P) generally taught theoretical courses, there was much effort required outside of class, and the student felt threatened by grades. At the opposite pole (i.e, second quadrant) were instructors who taught more applied courses, especially near the vertical axis. Moving around more toward the horizontal axis of this quadrant were those who did not require much effort outside of class and under whom the students did not feel threatened by grades.

Since the use of cluster analysis can be helpful, the K-means cluster analysis program was applied to the three dimensional solution supplied by KYST. A seven cluster solution was chosen and the cluster membership data appear in Table 2. The interviewers agreed that this grouping made some sense and supplied some characteristics of the groups.

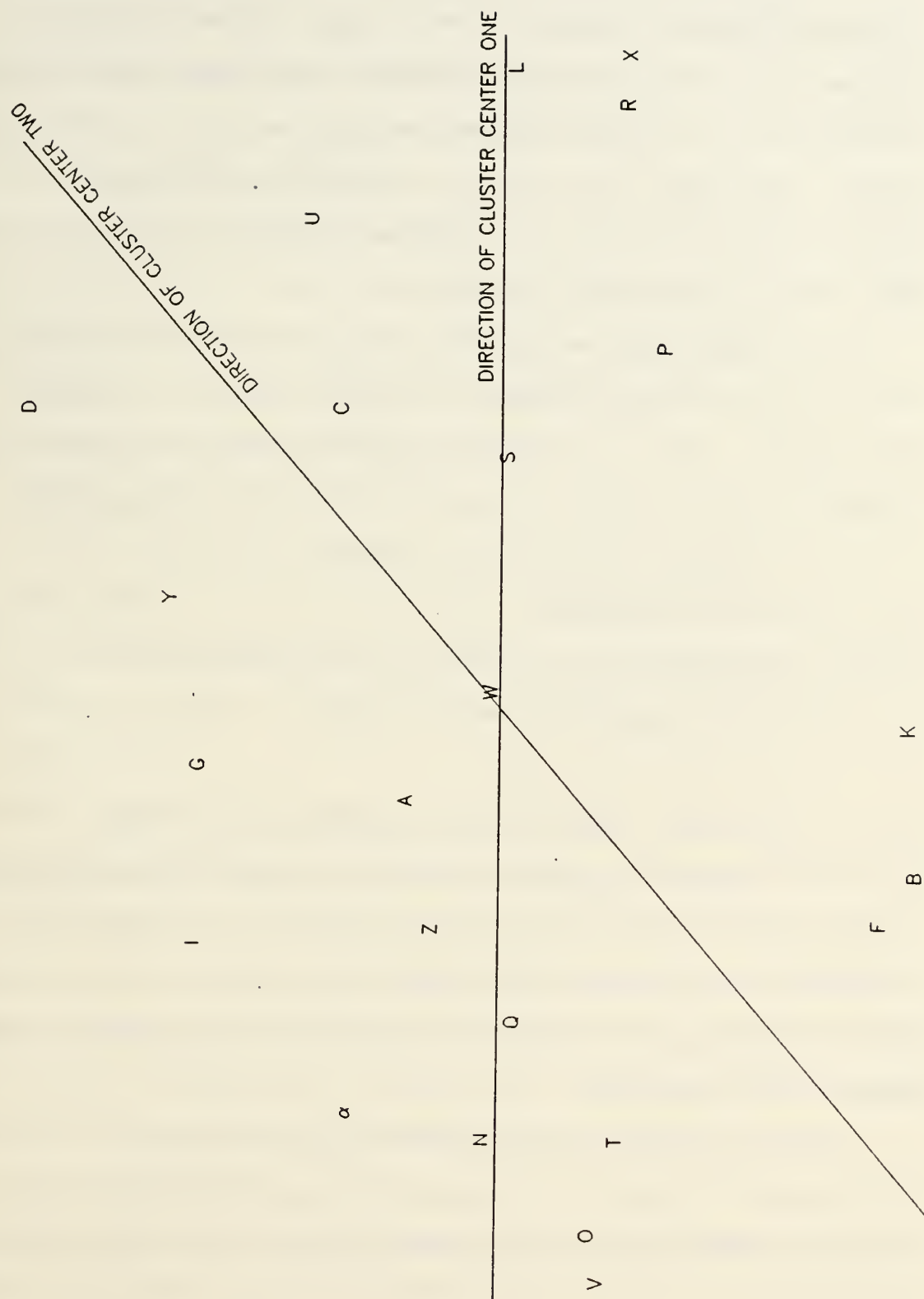
Table 2
Seven Clusters for the Three Dimensional
Solution of the September 1984 data.

	Members	Coordinates of Centers		
Cluster 1	I S R X	-.220	-.447	.028
Cluster 2	L G C	-.523	-.224	.134
Cluster 3	Y U	-.706	-.162	.407
Cluster 4	P W Z α A	-.057	.150	-.055
Cluster 5	N Q	.144	.584	-.254
Cluster 6	O K V T F B	.916	.117	-.102
Cluster 7	D	-1.540	.228	.061

It was not possible to supply additional graphics in time for the exit interview, but we will present an example of how the technique of cluster analysis can be used to help construct useful supplementary scatter plot projections. We have already selected a seven group clustering for the three dimensional MDS solution. Next the set of direction cosines for all possible pairs of cluster center vectors was computed from Table 2. From this one can discern that cluster centers one, three and four form a set of three clusters which is as non colinear as possible. They span the space and we can project the data onto the planes formed by any two of these center vectors selected from the three. Figure 6 was prepared to illustrate this idea. The directions of the cluster centers one and three provide an oblique coordinate system. It is hoped that this technique will reduce the amount of variability that is not contained in the planes of the scatter plot and provide a better display than that provided by the (default) principal components technique. Also, the cluster membership (table) should be useful in the exit interview.

FIGURE 6

SEPT 1984 DATA PROJECTED TO THE PLANE OF CLUSTERS ONE AND THREE



March 1985 Study.

The experience of the pilot study lead to a more organized and deeper effort for the next graduating class. For his master's thesis research, Lt J. McCourt developed user friendly software so that the respondents could read their instructions at an IBM 3278 terminal and enter their data directly into the machine. From there it was sent to a central file and processed. The MDS, cluster analysis and graphical output programs were executed and the exit interviews were held, this time involving 23 students.

In this experiment the students were asked to provide proximity information for all pairs instructors in terms of how close they were in teaching effectiveness. Also, on a scale of one to nine, they were asked to rate each instructor (or the course he taught) on each of the following bipolar scales:

0. Timeframe (recency) of the course.
1. Size of the class.
2. The applied vice theoretical nature of the course.
3. The anticipated severity of grading.
4. The pace of the course.
5. The effort required of the student outside of class.
6. The extent to which the course relied upon its prerequisites.

Items 2, 3, and 5 may be recognized as important characteristics that were identified in the September pilot study. The others were added by the author. Item 0 was subsequently deleted as it was stated in a confusing fashion and the responses were unreliable.

Additional bipolar scale information was made available in the form of retrieving the SOF data for the courses taken by these students. For sake of immediate reference these items (paraphrased and in the order of SOF number plus six) are:

7. Course organization.
8. Time in class spent effectively.
9. Instructor knew when student didn't understand the material.
10. Difficult concepts made understandable.
11. Confidence in Instructors knowledge of subject.
12. Felt free to ask questions.
13. Instructor was prepared for class.
14. The objectives were made clear.
15. Instructor made course a worthwhile learning experience.
16. Instructor stimulated interest in the subject area.
17. Instructor cared about student progress.
18. Overall rating of the instructor.
19. Overall rating of the course.

The scatter plots of proximities on fitted solutions appear in Figure 2 and the stress vs dimension information is in Table 1. Lt. McCourt favored studying the four dimensional solution. I'd like to draw attention to some features of the five dimensional solution.

A multiple regression was performed for each of the nineteen bipolar scales on the five dimensional MDS solution, rotated to principal components. The regression coefficients, normalized to be direction cosines in order to help identify the unknown factors, appear in Table 3 along with the squared multiple correlation coefficients. The most striking feature is that the multiple correlation for scale 11 (i.e. SOF item 5) is negligible. In other words, confidence in the instructor's knowledge of his subject is not an important variable for this class to discriminate among instructors.

In order to identify the important dimensions of the MDS solution space, we would prefer to find some bipolar scales with higher multiple correlation coefficients. Had a number of these been .9 or higher there would be great

confidence in identifying directions in the solution space that are dedicated to the corresponding scales, and the instructors could be scored on these scales. If several characteristics have the same direction then we can assert that the students perceive these items to be the same in terms of discriminating among teachers for their effectiveness.

It may not be possible for us to get high correlations simply because we are merging the perceptions of a large number of people. Different people can be expected to treat the value of pertinent characteristics in differing ways. Another interpretation of the presence of lower correlations is simply that we have not yet identified the discriminating characteristics correctly.

Let us turn to the question of identifying important directions in our five dimensional solution space. As a first step, we compute the direction cosines between all pairs of bipolar scales as represented by their regression coefficients. See Table B.4 in Appendix B. Study of these values reveals much structure.

Firstly, scales 3, 4, 5, and 6 all have about the same direction. Their submatrix of cosines is

	3.	4.	5.	6.
3. Stringency of grading	1.0	0.985	0.909	0.888
4. Pace of course	-	1.0	0.911	0.941
5. Outside effort	-	-	1.0	0.767
6. Rely on prerequisites	-	-	-	1.0

This general direction represents how onerous the course was for the student. Since "effort" - scale 5 - has the largest multiple correlation coefficient, let us use its direction to represent all of these.

Secondly, scale 2 is rather orthogonal to the four scales above. Its cosines with those scales are

	3.	4.	5.	6.
2. Applied vs theoretical	0.136	0.101	0.036	0.200

Table 3
Normalized Regression Coefficients
(Direction Cosines)

Multiple
Correlation
Coefficient

<u>Experimental</u>	<u>DIM1</u>	<u>DIM2</u>	<u>DIM3</u>	<u>DIM4</u>	<u>DIM5</u>	
Class size	0.2589	-0.0615	0.6034	-0.4174	-0.6250	.418
Applied vs theoretical	0.1920	-0.5889	-0.4226	0.3858	0.5372	.578
Grading policy	0.221	-0.5881	-0.3419	-0.0591	-0.6963	.541
Pace of course	0.1653	-0.4887	-0.4673	-0.0840	-0.7129	.694
Effort required outside class	0.0654	-0.5912	-0.3583	-0.4373	-0.5712	.774
Course relied upon prerequisites	0.0461	-0.3544	-0.6669	0.1185	-0.6429	.680
<u>SCALES</u>						
Course organization	-0.5916	-0.5587	0.4039	0.3970	0.1309	.498
Time in class spent effectively	-0.6857	-0.4763	0.2436	0.4937	0.0036	.516
Instructor knew when students didn't understand material	-0.6546	-0.0826	0.3558	0.1594	0.6424	.293
Difficult concepts made understandable	-0.5291	-0.0329	0.5403	0.5122	0.4059	.640
Confidence in instructors knowledge in subject	-0.2753	-0.0510	-0.3433	-0.3445	0.8277	.109
Felt free to ask questions	-0.5423	0.1217	0.2527	0.4140	0.6752	.516
Instructor prepared for class	-0.5332	-0.6405	0.5067	0.2080	0.0737	.478
Instructors objectives made clear	-0.4703	-0.8072	0.2965	0.1690	0.1038	.548
Instructors made course worthwhile learning experience	-0.6196	-0.0267	0.5665	-0.1371	0.4529	.525
Instructor stimulated interest in subject area	-0.6811	0.0024	0.7002	-0.2127	0.0265	.301
Instr. cared about student progress and did his share in helping to learn	-0.0589	-0.1654	0.5475	0.1054	0.5611	.622
Overall rating of instructor	-0.5929	0.5503	0.5429	0.1912	0.1197	.513
Overall rating of course	-0.2660	-0.7673	0.5241	-0.0856	-0.2419	.386

and, since applied vs theoretical is an important scale, (i.e. high multiple correlation), let us designate its direction in our five dimensional space. It does not appear to be identified with any of the other scales.

Thirdly, scales 7, 8, 14, and 13 form a cohesive set. Their set of direction cosines is

	7.	8.	14.	13.
7. Organization	1.0	0.966	0.929	0.970
8. Time spent effectively	-	1.0	0.862	0.897
14. Objectives clear	-	-	1.0	0.960
13. Prepared for class	-	-	-	1.0

This set represents organization in general, and scale 14 will be used to typify it, (multiple correlation = 0.548).

Fourthly, scales 10, 12, 17, and to a large degree 15 as well, form another coherent set having common direction. Their cosines are

	10.	12.	17.	15.
10. Difficult concepts	1.0	0.906	0.894	0.754
12. Felt free to question	-	1.0	0.858	0.691
17. Instructor cared	-	-	1.0	0.958
15. Worthwhile experience	-	-	-	1.0

This is an instructor-group interaction set and scale 10 will be used to represent it, (multiple correlation = 0.64).

Finally it is convenient to include scale 18, the overall rating of the instructor, to serve as a fifth direction to span our five dimensional solution. It is well correlated with other directions but no other scaled direction appears as being prominent and important. So let's include it. Now the direction cosine matrix of our five selected vectors of regression coefficients appears next:

	2.	5.	14.	10.	18.
2. Applied vs theoretical	1.0	0.037	0.375	0.100	0.113
5. Outside effort	-	1.0	0.212	-0.664	-0.057
14. Objectives clear	-	-	1.0	0.561	0.928
10. Difficult concepts	-	-	-	1.0	0.770
18. Overall rating	-	-	-	-	1.0

From the above table it is quite conspicuous that the overall rating is strongly correlated with organization (represented by "objectives made clear") and modestly well with instructor-group interaction (represented by "difficult concepts made understandable"). This is an interesting comment about this class. Note also that the other two dimensions are nearly orthogonal to the overall rating direction.

In a review of the scales that have been omitted, it may be seen that scale 1 (class size) has a modest shared direction (-0.667) with scale two (applied vs theoretical) but no noticeable communality with any of the other major scales. Its multiple correlation (0.418) is borderline among those used thus far. (The negative sign may be explained by the fact that the theoretical courses come early in the curriculum when the class sizes are large and the applied courses come later, generally are electives and have smaller class sizes.) Class size will be given no further consideration as a discriminator. The remaining scales (9, 11, 16, and 19) all have very small multiple correlation coefficients, and will be ignored.

The five scales identified above (applied vs theoretical, effort, organization, instructor-group interaction, and overall rating) may be used as a new basis for our five dimensional description of the student perception space. Suppose we are trying to make comparisons among instructors for administrative purposes. Suppose further that we do not want such comparisons to depend upon the first two of these scales.

We proceed to show how these may be removed. The plane of the directions of these two scales is spanned by these two vectors (i.e. the five component vectors for lines 2 and 5 in Table 3), and we can choose two orthogonal basis vectors in it. The rest is obtained by completing an orthonormal transformation. The subspace formed by the axes of these last three directions will be called the orthogonal complement of the applied vs theoretical - effort base plane. Once the data from our five dimensional solution have been rotated into the coordinates of this new basis, we need only study the last three components to achieve our goal of removing the effect of the first two scales.

When these last three axes have been rotated to their own principal components (call them R_1 , R_2 , R_3) it is interesting to note the variances and cumulated percentage variances of the data:

Subscale	R_1	R_2	R_3
Variance	0.4482	0.1994	0.0990
Cumulative percent of total	60	87	100

Also it is interesting to record how much change there is when scales 14, 10, and 18 are regressed on the subspace of R_1 , R_2 , and R_3 . The regression coefficients and the multiple correlation coefficients (not normalized this time) are:

Scale	Beta1	Beta2	Beta3	R-squared
14. Objectives clear	0.308	0.425	0.051	0.470
15. Difficult concepts	0.419	0.359	0.143	0.470
18. Overall rating	0.439	0.539	0.061	0.509

Attention is drawn to the following points:

The first two multiple correlations are down slightly from 0.548 and 0.640 resp., that were found in Table 3, but the last one has hardly changed - the original value being 0.513. The direction R_3 is hardly needed as far as these three scales are concerned. The overall rating favors R_2 slightly over R_1 as does the scale

representing organization. The scale representing instructor-student interaction has a small preference for R1 over R2. Thus a direction of instructor popularity could be constructed in the R1, R2 plane. Thus we have narrowed the scales of importance for this class and found a way to score the instructors on these scales.

In search of a better way to prepare graphical presentations for the exit interviews, the author chose to experiment with the following ad hoc technique. It involves the selection of sets of instructors whose data vectors have common direction. The data are to be projected on the planes of these directions. The goal is to reduce the degree of the third (and higher) dimensional variability when viewing vectors projected to planes. This provides an alternative to the formal cluster analysis technique applied to the September 1984 data.

Let us illustrate. The matrix of direction cosines for the sixteen objects in our orthogonal complement space are computed and appear in Table B.5. Upon scanning this table we extract the following submatrix.

INSTR.	C	1.					
	E	-.94	1.				
	K	-.97	.85	1.			
	M	-.99	.95	.98	1.		
	O	.97	-.94	.98	-.82	1.	
	P	.98	-.98	-.91	-.97	.99	1.
		C	E	K	M	O	P

These have about the same directions. The negative signs merely indicate the opposite pole of the same direction.

Setting these aside we look again for another separate set and come up with:

	A	1.			
	D	.95	1.		
INSTR.	I	.99	.96	1.	
	J	-.99	-.97	-1.	1.
	A		D	I	J

The cross matrix of direction cosines of these two groups will provide us with information about the plane (roughly) spanned by those two general directions

		INSTR.					
		C	E	K	M	O	P
	A	.32	-.41	-.45	-.47	.14	.29
INSTR.	D	.03	-.12	-.19	-.19	-.17	-.01
	I	.27	-.37	-.39	-.42	.09	.24
	J	-.24	.35	.36	.39	-.06	-.21

The largest magnitude is .47, which represents an angle of about 62 degrees, and the smallest is .01 (about 90 degrees). It is concluded that these two general directions are reasonably separate.

Further scanning in this fashion does not produce additional important groupings, so let us move on. To obtain firm directions we choose instructors C and J to represent their respective groups. They have the highest magnitude direction cosines in their respective sets. This done we project the data onto the plane of these two directions. The result appears in Figure 7.

This plot is useful in the following way. Instructor D is regarded as very different in style and this accounts for his isolation. This class reported that the three best instructors are E, G, and L and the two they held in low regard are O and B. These latter two are fairly isolated on this plot as is the pair G and L. Their positions only provide some general information about the interpretation of directions. Further interpretation can be discovered by asking the students what P and C have in common and asking the same for instructors A and I. Since the high quality instructor E is positioned not far from K, M and H, we may learn the teaching characteristics that these have in common.

In order to study the effect of the third dimension in this space let us consider two more plots. Figure 8 contains the data projected on the plane of direction C and the direction orthogonal to the plane presented in Figure 7, while Figure 9 contains the data projection to the plane of direction J and the same orthogonal. From the scale of this orthogonal axis we see that the data do not extend very deeply into the third direction. (It may have been better to plot each axis to a common scale.)

In Figure 8, instructors A and I are still close together but they have moved closer to the popular pair G and L. We begin to see some separation of C and P. Our very different instructor D is no longer isolated.

Figure 9 offers the best opportunity to project the instructors into a single popularity axis as the three favorite appear in the first quadrant and the poorly regarded ones appear in the third. Instructor H may be associated with some attractive characteristics; D is isolated again.

The act of drawing attention to features such as these during the exit interview enables the collection of organized information about qualities of instruction. It allows deeper understanding and over time, one can separate the idiosyncrasies of the various student groups from those qualities that have more permanence.

FIGURE 7 -- ORTHOGONAL COMPLEMENT SPACE
 MARCH 1985 DATA PROJECTED TO THE PLANE OF C AND J

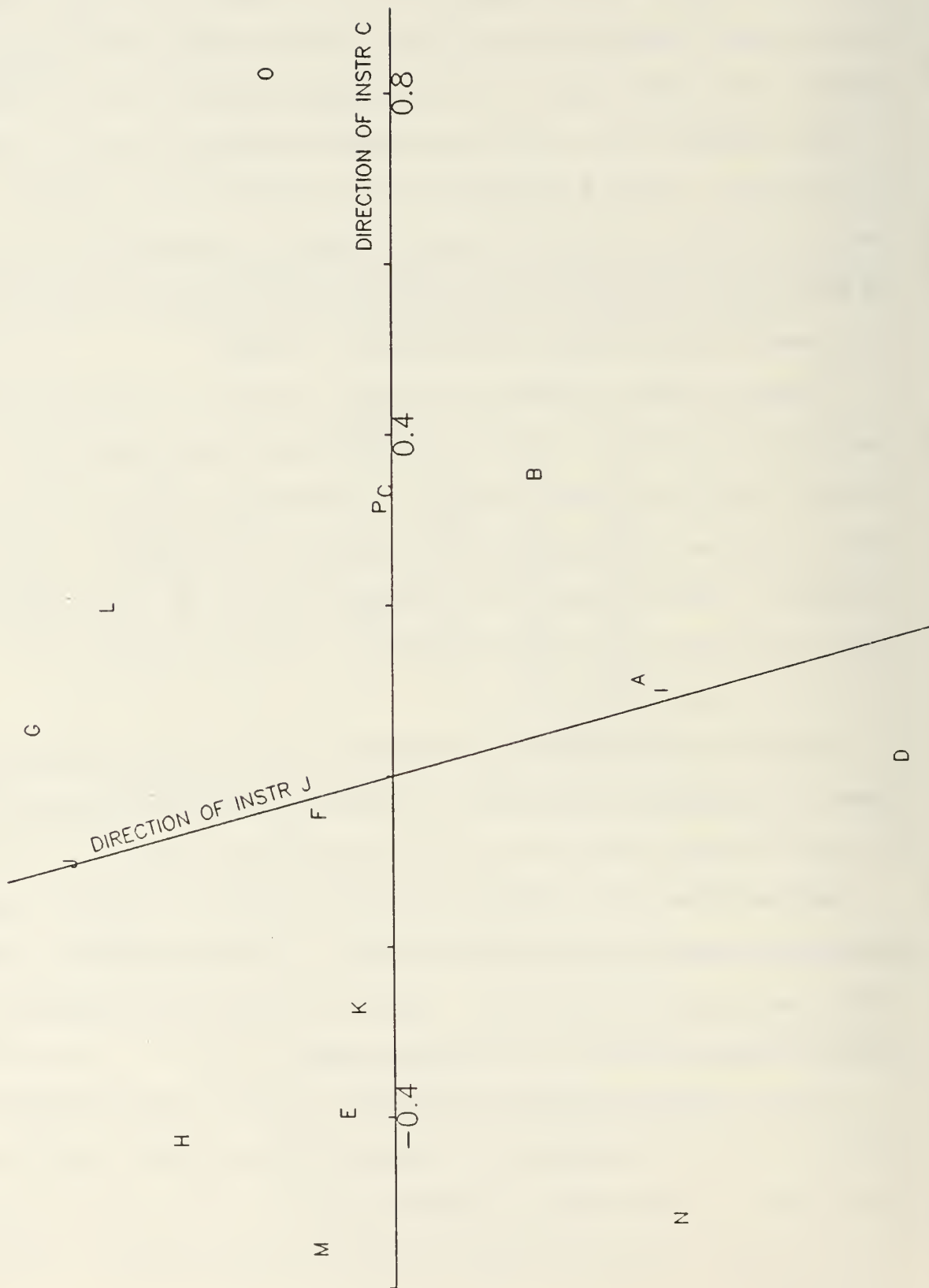


FIGURE 8 - ORTHOGONAL COMPLEMENT SPACE

MARCH 1985 DATA PROJECTED TO PLANE OF C AND NOT J

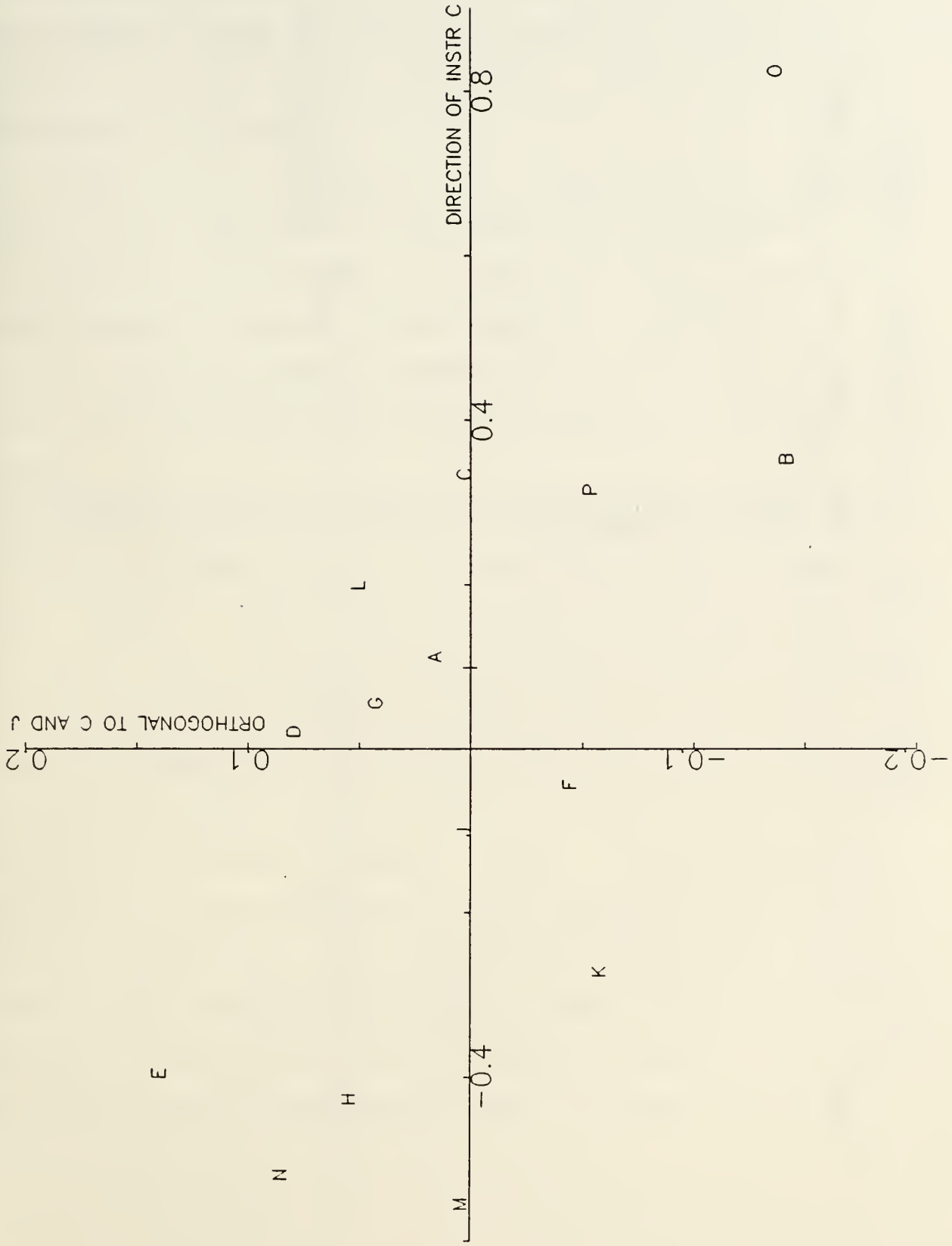
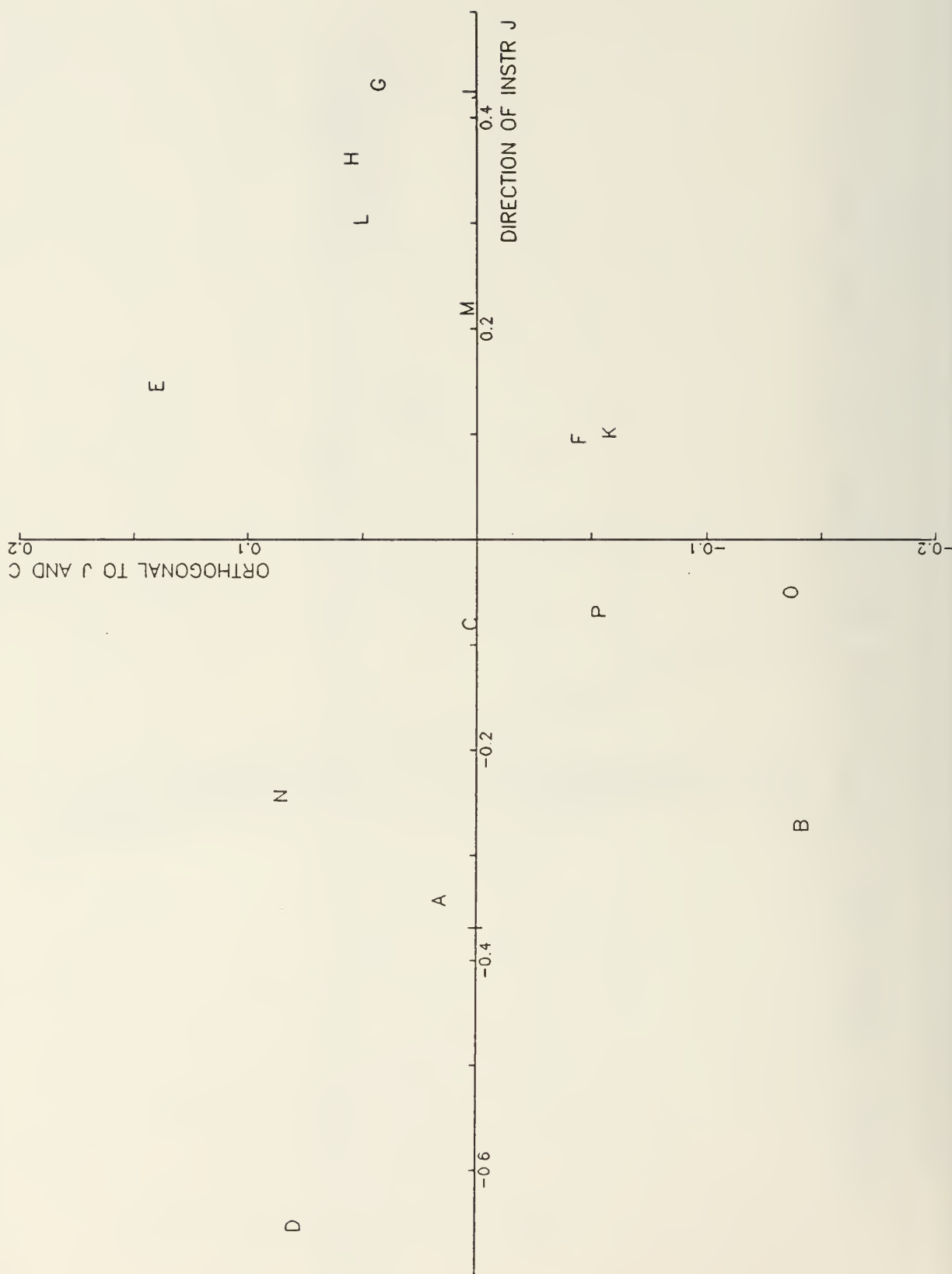


FIGURE 9 -- ORTHOGONAL COMPLEMENT SPACE
MARCH 1985 DATA PROJECTED TO THE PLANE OF J AND NOT C



V. Conclusions and Recommendations

MDS provides us with a valuable tool for gaining a deeper understanding of the student-group/course/instructor interface. The important effects can be discovered dynamically rather than having to be prescribed in advance. The necessity for cross classified data is given relief. Important effects may be separated without this requirement.

The technique allows us to recognize the individuality of the various classes of students. We can track trends in what students look for in teachers.

The developmental work exhibited so far suggests there may be present an unexplained dimension of teaching (or rather perception of instructor performance). On the other hand this vagueness may be due to the error introduced by the pooling of data from an entire student group prior to the structuring of an MDS solution.

For further development we recommend the use of individual scaling. That is, the conversion of proximity data to an MDS solution should be made for each individual respondent. Then the results can be pooled for the entire student group by applying linear (multivariate) scaling transformations tailored to each solution so that the group solution has as little variability as possible. In this way we would hope to gain more curvature in the stress versus dimension plot (so that the proper number of dimensions can be identified) and work with stress levels that are lower and more desirable. Moreover this approach should lead to reduced uncertainty in the interpretations gathered during the exit interviews.

The use of bipolar scales in the data collection activity should continue. They need to be well selected and not too many in number. Some of these should duplicate the SOF items. The original SOF data does not correlate as well with the MDS solution as does the data from the freshly collected bipolar scales.

Perhaps time alters the perception process. Lt. McCourt found that the overall rating of the instructor measured in the final quarter did not correlate well (50%) with retrieved scores from SOF Item 12.

Further recommendations include improvement of the user friendly programs. The maximum number of proximity values that can be supplied by the respondent who has had N instructors is $N(N-1)/2$, and it can be very time consuming to attempt to generate all of them. As there is much redundancy in these values it should be possible to specify a reasonable number in advance, say M , and present the respondent with M pairs of instructor names chosen at random from the much larger maximum value. To my knowledge this is a new feature of MDS and would require some study, planning, and program modification for its implementation.

Although the developmental work is not complete, we have already a potentially useful result. Referring to Figure 3, the first two principal components of the September 84 data, recall that instructors appearing in the third quadrant (or the courses they taught) were identified by the students as teaching theoretical (vice applied) courses, requiring much work outside of class, and having stringent grading policies. An examination of the SOF Item 12 scores awarded these instructors shows generally that they are low. This suggests that this particular class associates the identified characteristics with poor instruction.

It is possible to make immediate use of this information. Suppose, for example, that Instructor S were being considered for promotion or tenure. His SOF rating, although not high in the absolute sense, is quite high relative to this group of instructors having the common characteristics. Information of this type should be useful in the construction of the dossier.

REFERENCES

1. Burgess, C. T. Jr., and Vaughn, J. A., "A Multivariate Statistical Analysis of Student Opinion Questionnaires Concerning Teaching Effectiveness at the Naval Postgraduate School", Master's Thesis, U.S.N. Postgraduate School, September, 1973.
2. Kruskal, J. B. and Wish, M., Multidimensional Scaling, Sage University Paper No. 11, Series: Quantitative Applications in the Social Sciences, Beverly Hills, 1981.
3. Kruskal, J. B., Young, F. W., and Seery, J. B., How to Use KYST2A, A Very Flexible Program to do Multidimensional Scaling, Bell Telephone Laboratories, 1977.
4. McCourt, J. F., "Using Multidimensional Scaling to Describe Teacher Performance", Master's Theses, U.S.N. Postgraduate School, March 1985.
5. Read, R. R., and Zweig, H. J., "On the Quantification of Teacher Performance Using Student Opinion", NPS-55REZW72031A, U.S.N. Postgraduate School, March, 1972.
6. Read, R. R., "Components of Student-Faculty Evaluation Data", Educational and Psychological Measurement, 39, 1974.
7. SAS USER'S GUIDE: STATISTICS, SAS Institute Inc., 1982.

Appendix A: Raw Data

Sept. 1984

		Median Proximities																				
		W	O	N	Y	L	K	Z	V	I	U	G	T	S	F	R	α	D	C	Q	B	A
W	2.5																					
O	3.3	2.0																				
N	5.3	3.0	1.5																			
Y	6.7	5.0	8.3	4.8																		
L	1.0	-1.0	7.0	6.0	1.0																	
K	5.7	5.0	5.3	4.7	7.7	10.0																
Z	-1.0	.0	-1.0	1.0	.0	-1.0	4.5															
V	8.0	-1.0	2.0	2.0	8.0	9.0	.0	-1.0														
I	6.5	3.0	6.0	2.0	.8	-1.0	8.0	.0	-1.0													
U	5.3	4.0	6.3	5.3	3.0	2.0	7.3	-1.0	10.0	2.5												
G	8.5	4.0	8.5	4.2	.5	-1.0	7.3	.0	-1.0	2.3	3.0											
T	3.7	5.0	1.3	2.8	7.6	8.0	5.0	3.5	3.0	3.3	5.7	7.5										
S	1.0	3.7	3.7	4.5	5.6	2.0	4.3	4.0	5.0	4.7	4.3	5.7	3.3									
F	4.0	8.0	2.0	6.3	8.7	-1.0	5.7	8.0	-1.0	9.0	7.0	8.7	1.3	5.7								
R	1.0	3.0	10.0	7.0	4.4	1.0	4.0	7.5	9.0	7.5	2.0	5.5	6.8	4.0	5.5							
α	7.0	2.0	4.0	2.7	3.0	9.0	6.0	.5	1.0	2.0	5.0	1.8	5.7	5.1	7.7	6.2						
D	-1.0	-1.0	-1.0	-1.0	3.0	-1.0	10.0	-1.0	-1.0	5.0	-1.0	.0	10.0	5.0	-1.0	5.0	5.0	-1.0				
C	3.7	2.5	6.3	4.0	2.3	3.0	7.0	-1.0	7.0	2.0	1.7	2.0	7.3	5.0	9.0	2.0	4.0	-1.0				
Q	7.0	2.7	4.3	2.7	3.6	8.0	4.1	.0	1.0	4.3	5.3	2.7	5.3	5.7	7.7	6.0	2.2	5.0	4.0			
R	4.3	8.3	5.3	3.5	7.1	5.0	4.3	3.5	4.0	6.3	6.3	7.5	4.0	5.3	2.0	6.2	6.7	10.0	4.7	3.1		
A	-1.0	1.0	-1.0	2.7	2.3	-1.0	5.5	.0	-1.0	5.0	-1.0	2.3	5.3	4.3	6.0	5.0	2.5	5.0	-1.0	2.3	4.8	
X	-1.0	-1.0	-1.0	9.0	5.0	-1.0	4.0	10.0	-1.0	10.0	-1.0	5.0	10.0	5.0	10.0	1.5	5.5	-1.0	-1.0	5.0	9.0	7.0

Weights (Sample Size)

		W	O	N	Y	L	K	Z	V	I	U	G	T	S	F	R	α	D	C	Q	B	A
W	2																					
O	3	2																				
N	3	3	2																			
Y	3	3	3	6																		
L	1	0	1	1	1																	
K	3	0	4	3	0																	
Z	0	1	1	4	0																	
V	1	2	2	3	3																	
I	2	0	1	3	1																	
U	3	2	3	3	6																	
G	2	3	5	6	7																	
T	3	3	6	7	7																	
S	3	3	6	7	7																	
F	1	2	3	3	3																	
R	1	1	4	6	5																	
α	3	3	6	7	7																	
D	0	0	0	0	1																	
C	3	2	3	3	3																	
Q	3	3	6	7	7																	
B	3	3	6	7	7																	
A	3	3	6	7	7																	

[illegible][illegible]

Appendix B
Statistical Summaries

Table B.1

Sept. 1984 Data

Three Dimensional Solution Rotated to Principal Components

Instructor	1	2	3
P	0.387	-0.819	0.283
W	-0.218	0.060	-0.312
O	0.861	0.543	0.269
N	0.323	0.620	0.022
Y	-0.806	0.224	0.337
L	-0.185	-1.032	0.405
K	0.641	-0.336	-0.897
Z	-0.033	0.413	-0.043
V	0.500	0.764	-0.401
I	-0.357	0.636	0.510
U	-0.606	-0.548	0.478
G	-0.810	0.505	-0.006
T	1.060	0.295	0.575
S	0.218	-0.520	0.595
F	1.435	-0.302	0.239
R	-0.199	-0.997	-0.172
α	-0.199	0.820	-0.139
D	-1.539	0.228	0.061
C	-0.673	-0.204	0.004
Q	-0.034	0.549	-0.530
B	1.097	-0.263	-0.395
A	-0.222	0.274	-0.064
X	-0.641	-0.909	-0.819

Table B.2

March 1985 Data

Five Dimensional Solution Rotated to Principal Components

Instructor	1	2	3	4	5
A	-0.601	0.077	0.007	-0.694	-0.336
B	1.094	0.104	-0.55	-0.305	-0.055
C	-0.358	-0.076	-0.457	0.199	0.273
D	-0.558	-0.24	-0.061	-0.319	0.676
E	-0.754	-0.429	0.405	-0.202	0.151
F	0.648	0.25	0.086	0.265	0.181
G	-0.68	-0.128	-0.068	0.3	-0.349
H	0.427	-0.534	0.223	0.489	0.16
I	-0.285	0.681	0.258	-0.186	0.333
J	-0.234	0.29	0.311	0.087	-0.555
K	1.211	-0.308	0.052	-0.117	-0.224
L	-0.931	-0.252	-0.238	0.212	-0.292
M	0.755	0.331	0.812	0.158	0.01
N	-0.033	-0.345	0.558	-0.452	0.013
O	-0.141	1.186	-0.509	0.197	0.042
P	0.44	-0.606	-0.789	-0.036	-0.03

Simple Correlation Coefficients of the Instructors
Scores for All Pairs of Bipolar Scales, March 1985

41

Directed Cosines of the Angles Between All Pairs of Regression

Coefficient Vectors in table 3, March 1985

[illegible]

Direction Cosines Between All Pairs of Instructor Location Vectors
in the Three Dimensional Orthogonal Complement Space, March 1985

43

DISTRIBUTION LIST

	NO. OF COPIES
Library (Code 0142) Naval Postgraduate School Monterey, CA 93943-5100	4
Office of Research Administration (Code 012) Naval Postgraduate School Monterey, CA 93943-5100	1
Center for Naval Analyses 2000 Beauregard Street Alexandria, VA 22311	1
Operations Research Center, Room E40-164 Massachusetts Institute of Technology Attn: R. C. Larson and J. F. Shapiro Cambridge, MA 02139	1
Library (Code 55) Department of Operations Research Naval Postgraduate School Monterey, CA 93943-5100	1

217730

BF
39 BF
R4 39.5 Read
18 R42 Structured exit inter-
c 1985 views using MDS.
c.2

~~23~~¹⁶ JUL 85

DISPLAY

217730

BF
39.5 Read
R42 Structured exit inter-
1985 views using MDS.
c.2

genBF 39.5.R42 1985

Structured exit interviews using MDS /



3 2768 000 65969 2

DUDLEY KNOX LIBRARY